Dual-Core Processing Drives High-Performance Embedded Systems

Dual-core processors such as the AMD Opteron can overcome the problems associated with high-performance single-core CPUs, while delivering performance increases. Combined with improved microarchitecture, multithreading and HyperTransport connectivity, this technology is being harnessed to the needs of demanding embedded applications.

by Matt Stevenson and John Hill
WIN Enterprises

Dual-core CPUs have been commercially available since 2000 when IBM first introduced the IBM POWER4. They provide a method for gaining greater performance while avoiding the increases in form-factor, incremental heat and power requirements associated with the higher feature density of fast single-core processors. In pursuing dual-core CPUs, the major IC manufacturers have acknowledged that the historical approach of gaining performance by simply increasing CPU feature density has reached diminishing practical returns.

The current generation of high-performance CPUs (Table 1) is 90 nanometers (nm) between surface features, thus entering the realm of bona fide nanotechnology, which is 100 nm and below. However, at this extreme density, there are many unwanted effects. The industry has grown accustomed to ever improving performance with each CPU generation, but the current level of miniaturization of feature sizes is forcing IC manufacturers to look to more innovative solutions.

The problems caused by extreme feature density are interrelated. Electrical features in extreme proximity produce quantum effects, i.e., electrons that randomly tunnel across the CPU's features causing interference with normal signal transmission. At the highest frequencies, tunneling can become so extreme that it totally negates signal recognition.

Semiconductor Technology Generations
by Feature Size
(averaged across vendors)

Year	Feature size (in nanometers)	Comments
1982	1,500	
1993	600	
1998	250	
1999	180	
2001	130	
2003	**90**	**current generation**
2005	65	just beginning to appear

Table 1

To drive high performance across smaller, more powerful transistors requires more power. In turn, higher power results in unacceptable levels of waste heat as power (wattage) increases and produces more unwanted quantum effects. Machines with dense CPUs running at higher wattage are noisier, because they require additional, more powerful fans for cooling. Fan motors add yet more electrical noise.

Dual-core processors, such as the AMD Opteron, can mitigate these problems while at the same time enabling significant increases in performance.

The AMD Opteron Dual-Core Processor

The AMD Opteron processor is a high-density, 90-nm CPU, packing 233 million transistors on a 199 mm^2 die. The chip is microarchitected to lessen un-

wanted effects, principally through thread-level parallelism. It uses other technology, such as HyperTransport interconnect, in order to work smarter, not hotter.

Dual-core processors are most effective in applications that feature highly parallel processes. However, the technology can realize significant gains when applied to nearly any application that involves all but the simplest sequential number crunching. IBM, which has incorporated the dual-core AMD processor in some of its servers, reports 60% faster processing with a 2.2 GHz dual-core AMD Opteron processor versus AMD's 2.6 GHz single-core processor in tests using the Linpack HPL benchmark. Other tests, such as floating-point and integer processing, have yielded even better gains (Figure 1).

This increase in performance has generated interest among OEMs designing embedded systems for demanding, low-latency, high-performance applications, such as industrial automation, military, medical and security imaging, storage and telecommunications. The dual-core AMD Opteron processor enables basic reference designs that can be modified to meet these systems' needs for compactness, design longevity, lower power consumption, low latency and high reliability, often in harsh environments.

In response to market forces and evolving technology in x86 processors, such as the AMD Opteron and Pentium M, many designers of high-performance embedded systems are turning from highly specialized platforms to x86-based solutions. These systems typically run either Windows Embedded XP or Linux.

Regardless of the operating system chosen, it should be dual-core-aware in order to provide the benefit of multithreading. The dual-core AMD Opteron provides improved 32-bit legacy application support, in addition to concurrent 64-bit performance. This ability to support legacy applications enables a smooth upgrade path in the enterprise market and expands the dual-core Opteron's flexibility in the high-performance embedded market.

Terascala, which manufactures storage appliances for Linux-based clusters, is utilizing the dual-core Opteron CPU on

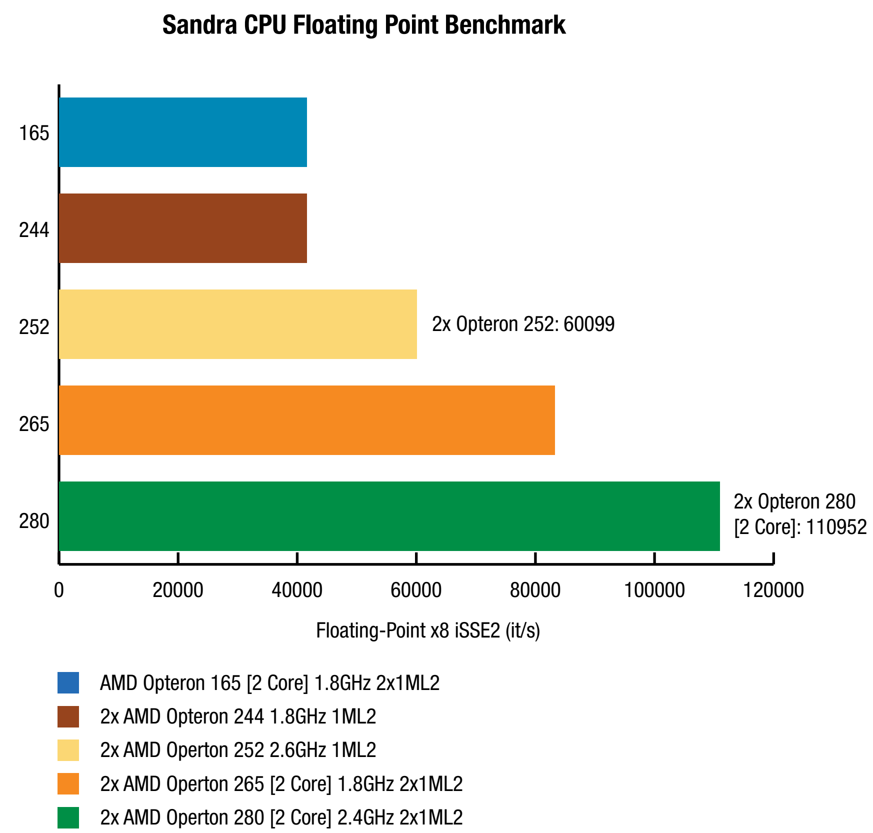

Figure 1 In floating-point performance tests (iterations per second) conducted at WIN Enterprises, the 2.4 GHz dual-core AMD Opteron 280 shows an 85% improvement over the 2.6 GHz single-core Opteron 252. This improvement can be attributed to the dual-core processor's ability to multithread its tasks.

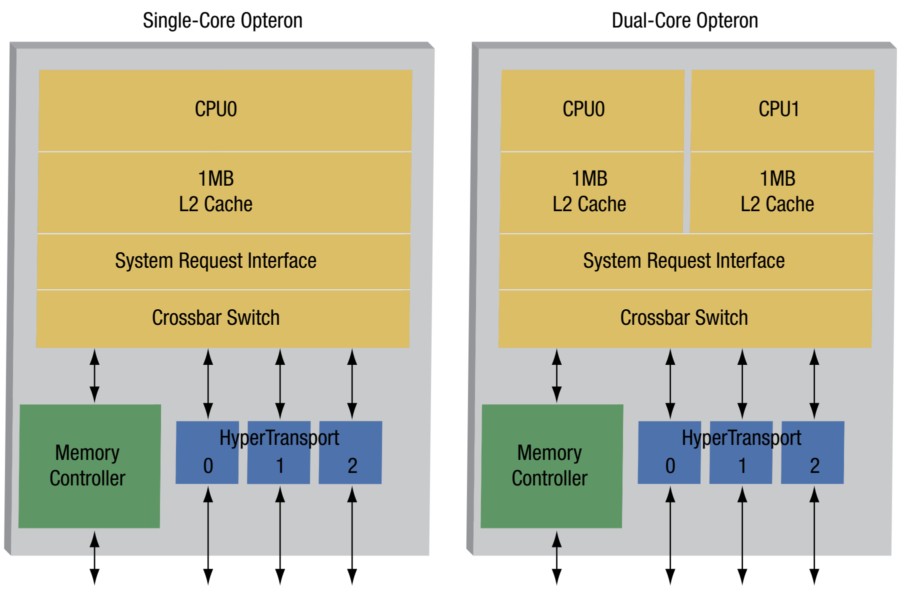

Figure 2 The dual-core AMD Opteron processor utilizes the same basic architecture as the single-core Opteron, but reduces board-level footprint. The two cores connect to a common crossbar that manages processing tasks and a dedicated L2 cache for each core provides scalability.

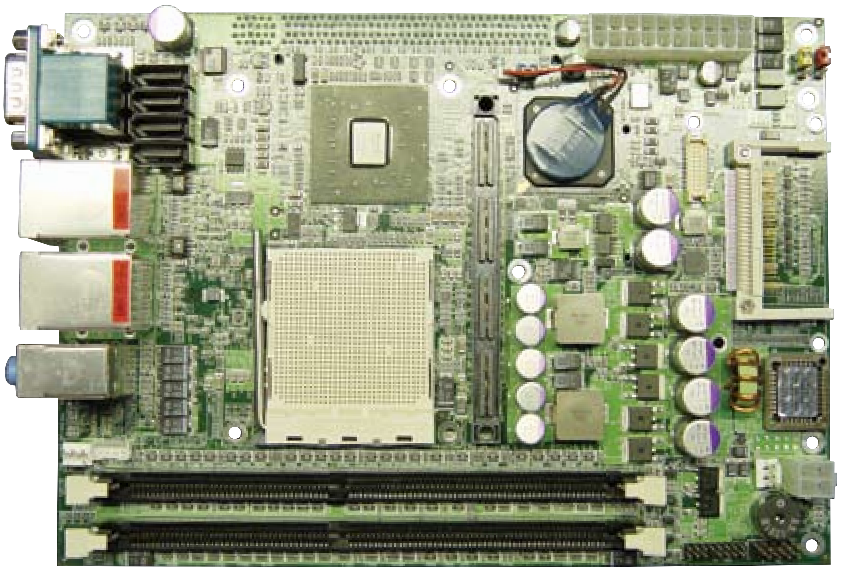

Figure 3 The MB-06047 EBX SBC from WIN Enterprises contains a low-power dual-core AMD Opteron CPU with a PCI Express slot, a CompactFlash socket, an ExpressCard socket, 4x SATA and a stackable HyperTransport connector.

motherboards co-designed and manufactured by WIN Enterprises. The combination of HyperTransport connectivity, improved microarchitecture and dual-core CPUs enables these storage systems to provide the high performance, scalability and high I/O throughput required by their enterprise customers.

Since Terascala rack-mounts several storage units into its cabinets, the benefits of a high-performance processor with a smaller footprint and less waste heat are especially important in serving data storage application needs. In addition, transaction-intensive storage environments require the dual-core architecture's low latency, which approaches real-time performance (Figure 2).

Multithreading

Multithreading separates programming into concurrent tasks across the two processing cores for enabling parallel processing. This results in more efficient processing and system resource utilization. An AMD Opteron dual-core design with two 2.2 GHz cores on a single die can outperform a 2.6 GHz single-core CPU because the dual cores can efficiently divide their processing tasks. This is true even though the clock frequency of the dual-core solution is slower than that of the single-core solution in order to control the dual-core CPU's level of waste heat production. However, even with multithreading techniques, higher performance is not a given.

The new x86 microarchitecture of the dual-core AMD Opteron processor is highly sophisticated. For instance, it features an integrated DDR memory controller. On-chip local memory in the form of L2 cache eliminates the need for the CPU to constantly fetch processing loads from RAM, as would be necessary with traditional Northbridge bus architectures.

Some of the Northbridge functionality, such as the memory controller, is designed into the CPU for greater throughput, resulting in a low-latency interconnect. This compares to the traditional Northbridge/Southbridge bus architecture, which can gate high system performance. This design innovation in microarchitecture is a major reason for the performance gains of both single- and dual-core AMD Opteron processors.

HyperTransport Technology

Originally begun at AMD, HyperTransport technology is a major advancement in chip-to-chip and board-to-board interconnection, and an important enabling technology in multicore CPUs, which is being applied to both commercial and embedded computing by several manufacturers.

A high-speed, low-latency technology, HyperTransport enables significant increases in communication speed between chips and I/O functions. It is scalable and can be used in expanding a dual-core design into a quad-core design through the use of stackable extension boards.

HyperTransport is both competitive with, and complementary to, PCI Express. Either can be used for both chip-to-chip and board-to-board interconnection, and they can be deployed either exclusively or together, depending on the application. A typical dual-core CPU design interfaces Hyper Transport with PCI Express. This takes advantage of PCI Express' support for a wide selection of chips as well as HyperTransport's throughput performance where it counts most, allowing embedded designs to be optimized for both function and performance.

HyperTransport features two unidirectional, point-to-point, high-speed connections that integrate chips, boards and other bus structures. The technology is also used in the integration of DDR memory with the CPU, enabling it to reside on the same die space. Separate HyperTransport links serve the I/O functions.

Applying Dual-Core Technology to High-Performance Embedded Designs

WIN Enterprises was one of the first board vendors to apply dual-core AMD Opteron technology to the needs of high-performance embedded computing. Working closely with AMD, the company had to innovate in order to solve several different problems in applying the technology to mobile servers, imaging devices and databank management applications.

First, an appropriate form-factor had to be decided upon. After evaluating a range of formats, including mini ITX, EBX, PICMG 1.3, ETX and EPIC, an

EBX SBC was selected (Figure 3), since it is increasingly sought by designers of high-performance embedded systems, partly for its small size.

WIN decided to populate the EBX form-factor of the MB-06047 SBC with state-of-the-art components. These included dual-core Opteron processors with HyperTransport, PCI Express, USB 2.0, ALC850 audio and Gigabit Ethernet.

In designing this board, the nVidia nForce 2200 chipset was chosen to work with the AMD CPU. However, the two had never been used together in a small form-factor, which presented some design challenges. The successful mating of dual-core and small form-factor was a breakthrough for the embedded OEM market.

Other challenges were overcome by designing a 10-layer motherboard rather than utilizing the traditional 6 layers.

Nextcom, a leading manufacturer of extreme performance, mobile, small-footprint computing products, is utilizing the Opteron dual-core CPUs on a related design, the MB-06048, which is a PICMG 1.3 form-factor. This is being used in a field-rugged, mobile data communications server used by military and government agencies. The advanced SBC enabled Nextcom to respond to market requirements for a distributed computing appliance that integrates legacy technology, performs multiple processes simultaneously, utilizes the advantages of COTS technology and allows application customization as market needs evolve.

The high-performance computing power of both single- and dual-core AMD Opteron processors is being leveraged in Nextcom's field-deployable units, the FleXtreme Vigor and NextDimension products. These small units top out at 2.6 GHz per processor. They use the stackable HyperTransport extension boards to offer quad-core CPU processing capability to military, government agency and other customers.

A WIN PICMG 1.3 reference design is also being used by a major workstation vendor in its medical imaging solution.

Software Considerations

Software is increasingly a concern in high-performance embedded designs, and that usually means Linux. To complement its efforts in high-performance small form-factor designs, WIN developed its own standard BIOS, as well as a downloadable Linux image for product testing and a Linux SDK.

Dual-core CPU technology is at the leading edge of high-performance embedded designs. Dual-core and HyperTransport technology enable a significant advance as a standardized x86-based platform that fulfills the requirements of low latency and high performance in a small form-factor. This approach is seeing a high level of OEM interest, evaluation and application.

WIN Enterprises
N. Andover, MA.
(978) 688-2000.
[www.win-ent.com].